



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 20, 2024 – 11:52 PM EDT

PDB ID : 2LEZ
BMRB ID : 17734
Title : Solution NMR structure of N-terminal domain of Salmonella effector protein PipB2. Northeast structural genomics consortium (NESG) target stt318a
Authors : Lemak, A.; Yee, A.; Houliston, S.; Garcia, M.; Daniels, C.; Savchenko, A.; Arrowsmith, C.; Montelione, G.T.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2011-06-27

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
wwPDB-ShiftChecker : v1.2
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

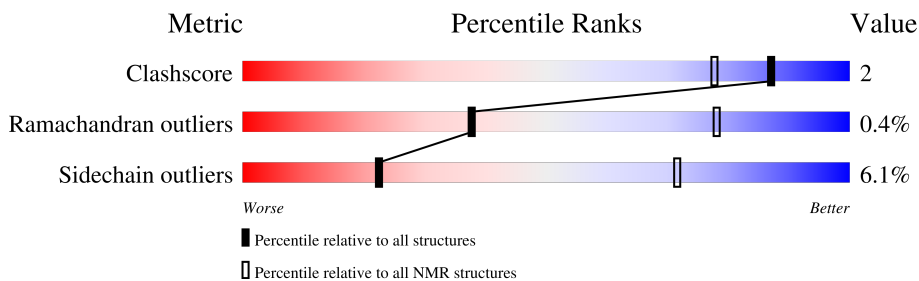
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 89%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	145	

2 Ensemble composition and analysis

This entry contains 20 models. Model 15 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:27-A:136 (110)	0.77	15

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 4 single-model clusters were found.

Cluster number	Models
1	5, 6, 8, 12, 13, 14, 15, 17, 18, 20
2	1, 3, 7, 9
3	2, 16
Single-model clusters	4; 10; 11; 19

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1863 atoms, of which 929 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Secreted effector protein pipB2.

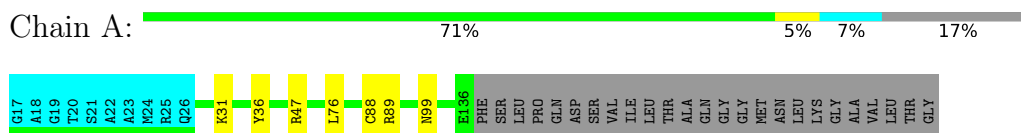
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	120	1863	577	929	166	183	8	0

4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

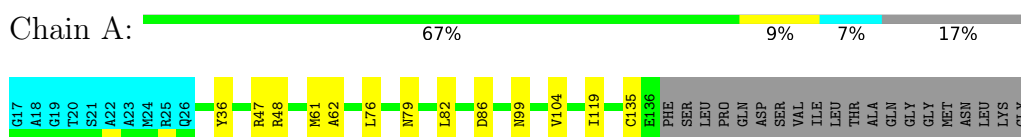
- Molecule 1: Secreted effector protein pipB2



4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 15. Colouring as in section 4.1 above.

- Molecule 1: Secreted effector protein pipB2



5 Refinement protocol and experimental data overview

The models were refined using the following method: *restrained molecular dynamics*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	structure solution	
CNS	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1644
Number of shifts mapped to atoms	1430
Number of unparsed shifts	0
Number of shifts with mapping errors	214
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	89%

6 Model quality [i](#)

6.1 Standard geometry [i](#)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	870	864	863	3±2
All	All	17400	17280	17260	69

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

5 of 27 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:88:CYS:SG	1:A:110:LYS:HD2	0.61	2.35	7	10
1:A:66:LYS:HD3	1:A:135:CYS:SG	0.58	2.38	19	1
1:A:61:MET:SD	1:A:82:LEU:HD11	0.58	2.38	4	6
1:A:62:ALA:HB1	1:A:135:CYS:SG	0.54	2.42	15	1
1:A:83:ASP:HA	1:A:89:ARG:HA	0.53	1.80	4	7

6.3 Torsion angles [i](#)

6.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	109/145 (75%)	100±2 (92±1%)	8±2 (7±2%)	0±1 (0±1%)	38 78
All	All	2180/2900 (75%)	2009 (92%)	163 (7%)	8 (0%)	38 78

All 5 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	111	GLY	3
1	A	48	ARG	2
1	A	70	PRO	1
1	A	44	GLY	1
1	A	46	ILE	1

6.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	97/121 (80%)	91±2 (94±2%)	6±2 (6±2%)	22 71
All	All	1940/2420 (80%)	1822 (94%)	118 (6%)	22 71

5 of 26 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	36	TYR	14
1	A	31	LYS	14
1	A	99	ASN	13
1	A	47	ARG	13
1	A	76	LEU	10

6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation i

The completeness of assignment taking into account all chemical shift lists is 89% for the well-defined parts and 88% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping i

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1644
Number of shifts mapped to atoms	1430
Number of unparsed shifts	0
Number of shifts with mapping errors	214
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 214) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	137	PHE	H	8.195	0.04	1
1	A	137	PHE	HA	4.608	0.04	1
1	A	137	PHE	HB2	3.144	0.04	2
1	A	137	PHE	HB3	2.9	0.04	2
1	A	137	PHE	HD1	7.19	0.04	3
1	A	137	PHE	HD2	7.19	0.04	3
1	A	137	PHE	HE1	7.249	0.04	3
1	A	137	PHE	HE2	7.249	0.04	3
1	A	137	PHE	C	175.421	0.40	1
1	A	137	PHE	CA	57.761	0.40	1
1	A	137	PHE	CB	39.817	0.40	1
1	A	137	PHE	CD1	131.876	0.40	3
1	A	137	PHE	CD2	131.876	0.40	3
1	A	137	PHE	CE1	131.384	0.40	3

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	137	PHE	CE2	131.384	0.40	3
1	A	137	PHE	N	120.627	0.40	1
1	A	138	SER	H	8.114	0.04	1
1	A	138	SER	HA	4.42	0.04	1
1	A	138	SER	HB2	3.789	0.04	2
1	A	138	SER	HB3	3.789	0.04	2
1	A	138	SER	C	173.781	0.40	1
1	A	138	SER	CA	58.158	0.40	1
1	A	138	SER	CB	64.051	0.40	1
1	A	138	SER	N	117.171	0.40	1
1	A	139	LEU	H	8.3	0.04	1
1	A	139	LEU	HA	4.597	0.04	1
1	A	139	LEU	HB2	1.569	0.04	2
1	A	139	LEU	HB3	1.537	0.04	2
1	A	139	LEU	HG	1.666	0.04	1
1	A	139	LEU	HD11	0.926	0.04	2
1	A	139	LEU	HD12	0.926	0.04	2
1	A	139	LEU	HD13	0.926	0.04	2
1	A	139	LEU	HD21	0.916	0.04	2
1	A	139	LEU	HD22	0.916	0.04	2
1	A	139	LEU	HD23	0.916	0.04	2
1	A	139	LEU	CA	53.207	0.40	1
1	A	139	LEU	CB	41.739	0.40	1
1	A	139	LEU	CG	27.193	0.40	1
1	A	139	LEU	CD1	25.409	0.40	2
1	A	139	LEU	CD2	23.665	0.40	2
1	A	139	LEU	N	124.815	0.40	1
1	A	140	PRO	HA	4.382	0.04	1
1	A	140	PRO	HB2	2.294	0.04	2
1	A	140	PRO	HB3	1.904	0.04	2
1	A	140	PRO	HG2	2.006	0.04	2
1	A	140	PRO	HG3	2.006	0.04	2
1	A	140	PRO	HD2	3.841	0.04	2
1	A	140	PRO	HD3	3.629	0.04	2
1	A	140	PRO	C	177.26	0.40	1
1	A	140	PRO	CA	63.449	0.40	1
1	A	140	PRO	CB	32.099	0.40	1
1	A	140	PRO	CG	27.643	0.40	1
1	A	140	PRO	CD	50.674	0.40	1
1	A	141	GLN	H	8.563	0.04	1
1	A	141	GLN	HA	4.261	0.04	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	141	GLN	HB2	2.076	0.04	2
1	A	141	GLN	HB3	1.969	0.04	2
1	A	141	GLN	HG2	2.361	0.04	2
1	A	141	GLN	HG3	2.361	0.04	2
1	A	141	GLN	HE21	6.865	0.04	2
1	A	141	GLN	HE22	7.528	0.04	2
1	A	141	GLN	C	175.913	0.40	1
1	A	141	GLN	CA	56.189	0.40	1
1	A	141	GLN	CB	29.468	0.40	1
1	A	141	GLN	CG	34.139	0.40	1
1	A	141	GLN	N	120.304	0.40	1
1	A	141	GLN	NE2	112.668	0.40	1
1	A	142	ASP	H	8.308	0.04	1
1	A	142	ASP	HA	4.58	0.04	1
1	A	142	ASP	HB2	2.683	0.04	2
1	A	142	ASP	HB3	2.64	0.04	2
1	A	142	ASP	C	176.242	0.40	1
1	A	142	ASP	CA	54.712	0.40	1
1	A	142	ASP	CB	41.227	0.40	1
1	A	142	ASP	N	120.882	0.40	1
1	A	143	SER	H	8.119	0.04	1
1	A	143	SER	HA	4.415	0.04	1
1	A	143	SER	HB2	3.822	0.04	2
1	A	143	SER	HB3	3.822	0.04	2
1	A	143	SER	C	174.409	0.40	1
1	A	143	SER	CA	58.705	0.40	1
1	A	143	SER	CB	63.942	0.40	1
1	A	143	SER	N	115.451	0.40	1
1	A	144	VAL	H	8.071	0.04	1
1	A	144	VAL	HA	4.078	0.04	1
1	A	144	VAL	HB	2.046	0.04	1
1	A	144	VAL	HG11	0.872	0.04	2
1	A	144	VAL	HG12	0.872	0.04	2
1	A	144	VAL	HG13	0.872	0.04	2
1	A	144	VAL	HG21	0.906	0.04	2
1	A	144	VAL	HG22	0.906	0.04	2
1	A	144	VAL	HG23	0.906	0.04	2
1	A	144	VAL	C	175.804	0.40	1
1	A	144	VAL	CA	62.533	0.40	1
1	A	144	VAL	CB	32.728	0.40	1
1	A	144	VAL	CG1	21.187	0.40	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	144	VAL	CG2	20.778	0.40	2
1	A	144	VAL	N	122.091	0.40	1
1	A	145	ILE	H	8.202	0.04	1
1	A	145	ILE	HA	4.131	0.04	1
1	A	145	ILE	HB	1.809	0.04	1
1	A	145	ILE	HG12	1.451	0.04	2
1	A	145	ILE	HG13	1.147	0.04	2
1	A	145	ILE	HG21	0.85	0.04	1
1	A	145	ILE	HG22	0.85	0.04	1
1	A	145	ILE	HG23	0.85	0.04	1
1	A	145	ILE	HD11	0.812	0.04	1
1	A	145	ILE	HD12	0.812	0.04	1
1	A	145	ILE	HD13	0.812	0.04	1
1	A	145	ILE	C	176.023	0.40	1
1	A	145	ILE	CA	61.07	0.40	1
1	A	145	ILE	CB	38.416	0.40	1
1	A	145	ILE	CG1	27.364	0.40	1
1	A	145	ILE	CG2	17.618	0.40	1
1	A	145	ILE	CD1	12.729	0.40	1
1	A	145	ILE	N	125.02	0.40	1
1	A	146	LEU	H	8.379	0.04	1
1	A	146	LEU	HA	4.441	0.04	1
1	A	146	LEU	HB2	1.655	0.04	2
1	A	146	LEU	HB3	1.572	0.04	2
1	A	146	LEU	HD11	0.9	0.04	2
1	A	146	LEU	HD12	0.9	0.04	2
1	A	146	LEU	HD13	0.9	0.04	2
1	A	146	LEU	HD21	0.833	0.04	2
1	A	146	LEU	HD22	0.833	0.04	2
1	A	146	LEU	HD23	0.833	0.04	2
1	A	146	LEU	C	177.253	0.40	1
1	A	146	LEU	CA	54.919	0.40	1
1	A	146	LEU	CB	42.551	0.40	1
1	A	146	LEU	CD1	25.191	0.40	2
1	A	146	LEU	CD2	23.448	0.40	2
1	A	146	LEU	N	127.046	0.40	1
1	A	147	THR	H	8.065	0.04	1
1	A	147	THR	HA	4.311	0.04	1
1	A	147	THR	HB	4.235	0.04	1
1	A	147	THR	HG21	1.173	0.04	1
1	A	147	THR	HG22	1.173	0.04	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	147	THR	HG23	1.173	0.04	1
1	A	147	THR	C	173.261	0.40	1
1	A	147	THR	CA	61.59	0.40	1
1	A	147	THR	CB	70.066	0.40	1
1	A	147	THR	CG2	21.513	0.40	1
1	A	147	THR	N	115.077	0.40	1
1	A	148	ALA	H	7.926	0.04	1
1	A	148	ALA	HA	4.109	0.04	1
1	A	148	ALA	HB1	1.313	0.04	1
1	A	148	ALA	HB2	1.313	0.04	1
1	A	148	ALA	HB3	1.313	0.04	1
1	A	148	ALA	CA	53.984	0.40	1
1	A	148	ALA	CB	20.274	0.40	1
1	A	148	ALA	N	131.881	0.40	1
1	A	149	GLN	HA	4.108	0.04	1
1	A	149	GLN	HB2	2.077	0.04	2
1	A	149	GLN	HB3	2.077	0.04	2
1	A	149	GLN	HG2	2.297	0.04	2
1	A	149	GLN	HG3	2.297	0.04	2
1	A	149	GLN	CA	57.437	0.40	1
1	A	149	GLN	CB	30.714	0.40	1
1	A	149	GLN	CG	34.323	0.40	1
1	A	151	GLY	HA2	3.962	0.04	2
1	A	151	GLY	HA3	3.962	0.04	2
1	A	151	GLY	C	176.5	0.40	1
1	A	151	GLY	CA	45.539	0.40	1
1	A	152	MET	H	8.191	0.04	1
1	A	152	MET	HA	4.293	0.04	1
1	A	152	MET	C	178.402	0.40	1
1	A	152	MET	CA	59.772	0.40	1
1	A	152	MET	N	119.418	0.40	1
1	A	153	ASN	H	7.357	0.04	1
1	A	153	ASN	N	119.665	0.40	1
1	A	154	LEU	HA	4.452	0.04	1
1	A	154	LEU	HB2	1.653	0.04	2
1	A	154	LEU	HB3	1.575	0.04	2
1	A	154	LEU	HG	1.602	0.04	1
1	A	154	LEU	HD11	0.895	0.04	2
1	A	154	LEU	HD12	0.895	0.04	2
1	A	154	LEU	HD13	0.895	0.04	2
1	A	154	LEU	HD21	0.832	0.04	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	154	LEU	HD22	0.832	0.04	2
1	A	154	LEU	HD23	0.832	0.04	2
1	A	154	LEU	C	176.515	0.40	1
1	A	154	LEU	CA	55.109	0.40	1
1	A	154	LEU	CB	42.558	0.40	1
1	A	154	LEU	CG	26.873	0.40	1
1	A	154	LEU	CD1	25.422	0.40	2
1	A	154	LEU	CD2	23.625	0.40	2
1	A	155	LYS	H	7.684	0.04	1
1	A	155	LYS	N	120.201	0.40	1
1	A	158	VAL	HA	4.029	0.04	1
1	A	158	VAL	HB	2.061	0.04	1
1	A	158	VAL	HG11	0.879	0.04	2
1	A	158	VAL	HG12	0.879	0.04	2
1	A	158	VAL	HG13	0.879	0.04	2
1	A	158	VAL	HG21	0.861	0.04	2
1	A	158	VAL	HG22	0.861	0.04	2
1	A	158	VAL	HG23	0.861	0.04	2
1	A	158	VAL	CA	63.749	0.40	1
1	A	158	VAL	CB	33.159	0.40	1
1	A	158	VAL	CG1	21.718	0.40	2
1	A	158	VAL	CG2	20.124	0.40	2
1	A	160	THR	HA	4.376	0.04	1
1	A	160	THR	HB	4.265	0.04	1
1	A	160	THR	HG21	1.189	0.04	1
1	A	160	THR	HG22	1.189	0.04	1
1	A	160	THR	HG23	1.189	0.04	1
1	A	160	THR	C	174.2	0.40	1
1	A	160	THR	CA	61.917	0.40	1
1	A	160	THR	CB	70.079	0.40	1
1	A	160	THR	CG2	21.731	0.40	1
1	A	161	GLY	H	7.973	0.04	1
1	A	161	GLY	HA2	3.761	0.04	2
1	A	161	GLY	HA3	3.761	0.04	2
1	A	161	GLY	CA	46.316	0.40	1
1	A	161	GLY	N	117.256	0.40	1

7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	138	-0.41 ± 0.07	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	126	0.02 ± 0.10	None needed (< 0.5 ppm)
$^{13}\text{C}'$	120	-0.09 ± 0.06	None needed (< 0.5 ppm)
^{15}N	121	0.09 ± 0.59	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 89%, i.e. 1341 atoms were assigned a chemical shift out of a possible 1512. 0 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	522/545 (96%)	216/221 (98%)	207/220 (94%)	99/104 (95%)
Sidechain	753/891 (85%)	514/575 (89%)	230/276 (83%)	9/40 (22%)
Aromatic	66/76 (87%)	33/37 (89%)	33/37 (89%)	0/2 (0%)
Overall	1341/1512 (89%)	763/833 (92%)	470/533 (88%)	108/146 (74%)

7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:

